

Achieving the Uniform Rate Region of General Multiple Access Channels by Polar Coding

Hessam MahdaviFar, Mostafa El-Khamy, Jungwon Lee, Inyup Kang

Abstract—We consider the problem of polar coding for transmission over m -user multiple access channels. In the proposed scheme, all users encode their messages using a polar encoder, while a joint successive cancellation decoder is deployed at the receiver. The encoding is done separately across the users and is independent of the target achievable rate, in the sense that the encoder core is the regular Arıkan’s polarization matrix. For the code construction, the positions of information bits and frozen bits for each of the users are decided jointly. This is done by treating the whole polar transformation across all the m users as a single polar transformation with a certain *base code*. We prove that the *covering radius* of the dominant face of the uniform rate region is upper bounded by $r = \frac{(m-1)\sqrt{m}}{L}$, where L represents the length of the base code. We then prove that the proposed polar coding scheme achieves the whole uniform rate region, with small enough resolution characterized by r , by changing the decoding order in the joint successive cancellation decoder. The encoding and decoding complexities are $O(N \log N)$, where N is the code block length, and the asymptotic block error probability of $O(2^{-N^{0.5-\epsilon}})$ is guaranteed. Examples of achievable rates for the case of 3-user multiple access channel are provided.

Index Terms—polar code, multiple access channel, uniform rate region

I. INTRODUCTION

POLAR codes were introduced by Arıkan in the seminal work of [1]. They are the first family of codes for the class of binary-input symmetric discrete memoryless channels that are provable to be capacity-achieving with low encoding and decoding complexity. Construction of polar codes is based on a phenomenon called the *channel polarization*. Arıkan proves that as the block length goes to infinity the channels seen by individual bits through a certain transformation called the *polar transformation* start polarizing which means that they approach either a noise-less channel or a pure-noise channel. In general, the constructed polar codes, together with the low complex *successive cancellation decoder*, achieve the *symmetric capacity* of all binary-input memoryless channels, where the symmetric capacity of a binary-input channel is the mutual information between the input and output of the channel assuming that the input distribution is uniform.

Polar codes and polarization phenomenon have been successfully applied to various problems such as wiretap channels [2], data compression [3], [4] and multiple access channels [5], [6]. The notion of channel polarization has been extended to two user multiple access channels (MAC) [5] and later to m -user MAC [6], wherein a technique is described to polarize a given binary-input MAC same as in Arıkan’s groundbreaking work of

[1]. The authors convert multiple uses of this MAC into single uses of extremal MACs and characterize the set of all extremal MACs in the asymptotic sense.

The capacity region of multiple access channels is fully characterized by Ahlswede [7] and Liao [8] for the case that the sources transmit independent messages. However, in this paper, we are only interested in the *uniform rate region*. For a multiple access channel, the uniform rate region is the achievable region corresponding to the case that the input distributions are uniform. The single user counterpart of uniform rate region is indeed symmetric capacity. It has been shown that at least one point on the *dominant face* of the uniform rate region can be achieved by the polar code constructed based on MAC polarization [5], [6]. But the problem of achieving the entire uniform rate region for the general case of m -user MAC remained open.

The core idea behind the approach of this paper, which distinguishes it from the approach taken in the prior works of [5], [6] can be explained as follows. It is well-known that the *corner points* of the uniform rate region are achievable with capacity-achieving single user codes. We observe that the successive method of decoding transmitted messages of different users at corner points follows the same logic as in the successive cancellation decoder of polar codes. Therefore, if polar encoders are deployed by the users, then the receiver can implement a joint successive cancellation decoder in such a way that all the messages are decoded in a unified manner. It can be shown that any particular decoding order will result in a certain achievable set of rates. Then in order to achieve different rates, the idea is to shuffle not only the encoded messages but also the encoded bits in different messages. The decoding order will be determined a priori and will be known at both the transmitters and the receiver. We show the direct relation between the decoding order and the polarization base code which guarantees the channel polarization. Furthermore, we show how to change the decoding order to approach all the points in the uniform rate region using the single user channel polarization theorem.

In another related work, Arıkan proposed a scheme for the Slepian-Wolf source coding problem, which is the dual of the special case of two-user MAC, based on *monotone* chain rule expansions and it is shown that the uniform rate region is achievable with polar coding [9]. In [10], [11], we showed how the polar coding with joint successive cancellation decoding can be used to achieve the uniform rate region of two-user multiple access channels, by regarding the MAC as one level of polarization, along with methods to improve the finite length performance. A method for improving the performance of two-user MAC polar coding with list decoding has been described in [12]. A straightforward generalization of these methods for

Hessam MahdaviFar, Mostafa El-Khamy, Junwon Lee and Inyup Kang are with Modem R&D, Samsung Electronics, San Diego, CA 92121, U.S.A. (e-mail: {h.mahdaviFar, mostafa.e, jungwon2.lee, inyup.kang}@samsung.com).

the m -user case provides achievability of the one-dimensional edges of the dominant face region. However, the dominant face of the uniform rate region is in general a polytope in an $m - 1$ -dimensional plane in the m -dimensional space and the main question was whether one can approach all points on the dominant face of the uniform rate region by changing the decoding order, thereby establishing the achievability of the entire region.

As opposed to the approaches taken in [9], [11], [12], we do not limit ourselves to a certain set of polarization base codes. In this paper, the set of all possible permutations are considered for the decoding order, assuming that the decoding order within the input bits of each user is preserved. Any polarization base code of length L leads to an m -tuple rate, which is proved to be achievable with polar coding. We further prove that the set of these achievable m -tuple rates cover the entire dominant face assuming m -dimensional balls of radius $\frac{(m-1)\sqrt{m}}{L}$ centered around them. This shows that all the points on the dominant face of the uniform rate region can be approached as the code block length and the length of polarization base codes grow large.

The rest of this paper is organized as follows. In Section II, we review some background on polar codes and multiple access channels. In Section III, the special case of 3-user MAC is considered and the set of all achievable rates for base codes of length 2 are derived for an example. In Section IV, we introduce the general MAC polarization base codes and prove that their corresponding achievable rates cover the entire dominant face of the uniform rate region. In Section V, we discuss how general MAC polar transformations can be built upon polarization base codes and establish the channel polarization theory. We also discuss the decoding method and provide comparison with prior work on MAC polar coding. At the end, we conclude the paper in Section VI.

II. PRELIMINARIES

A. Polar codes

In this subsection, we provide a brief overview of the groundbreaking work of Arikan [1] and others [13]–[15] on polar codes and channel polarization.

Polar codes are constructed based upon a phenomenon called *channel polarization* discovered by Arikan [1]. The basic polarization matrix is given as

$$G = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad (1)$$

Consider two independent copies of a binary-input discrete memoryless channel (B-DMC) $W : \{0, 1\} \rightarrow \mathcal{Y}$. The two input bits (u_1, u_2) , drawn from independent uniform distributions, are multiplied by G and then transmitted over the two copies of W . One level of channel polarization is the mapping $(W, W) \rightarrow (W^-, W^+)$, where $W^- : \{0, 1\} \rightarrow \mathcal{Y}^2$, and $W^+ : \{0, 1\} \rightarrow \{0, 1\} \times \mathcal{Y}^2$ with the following channel transformation

$$\begin{aligned} W \boxtimes W(y_1, y_2 | u_1) &= \frac{1}{2} \sum_{u_2 \in \{0, 1\}} W(y_1 | u_1 \oplus u_2) W(y_2 | u_2), \\ W \circledast W(y_1, y_2, u_1 | u_2) &= \frac{1}{2} W(y_1 | u_1 \oplus u_2) W(y_2 | u_2). \end{aligned} \quad (2)$$

$W \boxtimes W$ and $W \circledast W$ are also denoted by W^+ and W^- . The *bit-channels* W^- and W^+ are indeed the channels that u_1 and u_2 observe assuming the following scenario: the first bit u_1 is decoded assuming u_2 is noise, then u_2 is decoded assuming that u_1 is decoded successfully and is known.

The channel polarization is continued recursively by further splitting W^- and W^+ to get W^{--} , W^{-+} , W^{+-} , W^{++} etc. This process can be explained best by means of Kronecker powers of G . The Kronecker powers of G are defined by induction. Let $G^{\otimes 1} = G$ and for any $n > 1$:

$$G^{\otimes(n)} = \begin{bmatrix} G^{\otimes(n-1)} & 0 \\ G^{\otimes(n-1)} & G^{\otimes(n-1)} \end{bmatrix}$$

It can be observed that $G^{\otimes(n)}$ is a $2^n \times 2^n$ matrix. Let $N = 2^n$. Then $G^{\otimes n}$ is the $N \times N$ polarization matrix. Let (U_1, U_2, \dots, U_N) , denoted by U_1^N , be a block of N independent and uniform binary random variables. The polarization matrix $G^{\otimes n}$ is applied to U_1^N to get $X_1^N = U_1^N G^{\otimes n}$. Then X_i 's are transmitted through N independent copies of a binary-input discrete memoryless channel (B-DMC) W . The output is denoted by Y_1^N . This transformation with input U_1^N and output Y_1^N is called the polar transformation. In this transformation, N independent uses of W is transformed into N bit-channels, described next. Following the convention, random variables are denoted by capital letters and their instances are denoted by small letters. Let $W^N : \mathcal{X}^N \rightarrow \mathcal{Y}^N$ denote the channel consisting of N independent copies of W i.e.

$$W^N(y_1^N | x_1^N) \stackrel{\text{def}}{=} \prod_{i=1}^N W(y_i | x_i) \quad (3)$$

The combined channel \widetilde{W} is defined with transition probabilities given by

$$\widetilde{W}(y_1^N | u_1^N) \stackrel{\text{def}}{=} W^N(y_1^N | u_1^N G^{\otimes n}) \quad (4)$$

For $i = 1, 2, \dots, N$, the bit-channel $W_N^{(i)}$ is defined as follows:

$$W_N^{(i)}(y_1^N, u_1^{i-1} | u_i) \stackrel{\text{def}}{=} \frac{1}{2^{n-1}} \sum_{u_{i+1}^N \in \{0, 1\}^{n-i}} \widetilde{W}(y_1^N | u_1^N) \quad (5)$$

Intuitively, this is the channel that bit u_i observes through a *successive cancellation decoder*, deployed at the output. Under this decoding method, proposed by Arikan for polar codes [1], all the bits u_1^{i-1} are already decoded and are assumed to be available at the time that u_i is being decoded. The channel polarization theorem states that as N goes to infinity, the bit-channels start polarizing meaning that they either become a noise-less channel or a pure-noise channel.

In order to measure how good a binary-input channel W is, Arikan uses the *Bhattacharyya parameter* of W , denoted by $Z(W)$ [1], defined as

$$Z(W) \stackrel{\text{def}}{=} \sum_{y \in \mathcal{Y}} \sqrt{W(y|0)W(y|1)}$$

It is easy to show that the Bhattacharyya parameter $Z(W)$ is always between 0 and 1. Channels with $Z(W)$ close to zero are almost noiseless, while channels with $Z(W)$ close to one are almost pure-noise channels. More precisely, it can be proved

that the probability of error of a binary symmetric memoryless channel (BSM) is upperbounded by its Bhattacharyya parameter. Let $[N]$ denotes the set of positive integers less than or equal to N . The set of *good bit-channels* $\mathcal{G}_N(W, \beta)$ is defined for any $\beta < \frac{1}{2}$ [13], [14]:

$$\mathcal{G}_N(W, \beta) \stackrel{\text{def}}{=} \left\{ i \in [N] : Z(W_N^{(i)}) < 2^{-N^\beta/N} \right\} \quad (6)$$

Then the channel polarization theorem is proved by showing that the fraction of good bit-channels approaches the symmetric capacity $I(W)$, as N goes to infinity [13]. This theorem readily leads to a construction of capacity-achieving *polar codes*. The idea is to transmit the information bits over the good bit-channels while freezing the input to the other bit-channels to a priori known values, say zeros. The decoder for this constructed code is the successive cancellation decoder of Arkan [1], where it is further proved that the frame error probability under successive cancellation decoding is upper bounded by the sum of Bhattacharyya parameters of the selected good bit-channels, which is 2^{-N^β} by the particular choice of good bit-channels in (6).

B. Multiple access channel

Let $W : \mathcal{X}^m \rightarrow \mathcal{Y}$ be an m -user MAC, where \mathcal{X} is the binary alphabet and \mathcal{Y} is the output alphabet. With slight abuse of notation, the channel is described by the transition probability $W(y|x[1], x[2], \dots, x[m])$ for any $x[j] \in \mathcal{X}$, $j \in [m]$, and $y \in \mathcal{Y}$. Namely, $W(y|x[1], x[2], \dots, x[m])$ denote the probability of receiving y given that $x[j]$ is transmitted by the j -th user user, for $j \in [m]$. For any $\mathcal{J} \subset [m]$, let $x[\mathcal{J}]$ denote the set $\{x[j] : j \in \mathcal{J}\}$.

Let $X[1], X[2], \dots, X[m]$ be independent and uniform binary random variables generated by users 1, 2, \dots , m . Then the uniform rate region of W , denoted by $\mathcal{S}(W)$, is defined to be the set of all m -tuples $\mathbf{R} = (R_1, R_2, \dots, R_m) \in \mathbb{R}^m$ such that

$$0 \leq \sum_{j \in \mathcal{J}} R_j \leq I(X[\mathcal{J}]; Y, X[\mathcal{J}^c]), \forall \mathcal{J} \subseteq [m]. \quad (7)$$

The uniform rate region is the set of all achievable m -tuples of rates assuming that the input distributions are uniform. Let $I(W)$, which is also called the *uniform sum-rate* of W , be defined as follows:

$$I(W) = I(X[1], X[2], \dots, X[m]; Y)$$

In general, in the context of this paper, any point $\mathbf{R} = (R_1, R_2, \dots, R_m) \in \mathbb{R}^m$ is regarded as an m -tuple of rates and hence the $\sum R_j$, is referred to as the sum-rate of \mathbf{R} .

The dominant face of the uniform rate region, denoted by $\mathcal{D}(W)$, is defined to be the set of points in $\mathcal{S}(W)$, with the maximum sum-rate $I(W)$ i.e. the right inequality of (7) is in fact equality for $\mathcal{J} = [m]$. It can be observed that, the achievability of the uniform rate region $\mathcal{S}(W)$ is equivalent to the achievability of its dominant face $\mathcal{D}(W)$. Therefore, our focus throughout this paper is on the achievability of $\mathcal{D}(W)$.

III. THE PROPOSED SCHEME FOR 3-USER MULTIPLE ACCESS CHANNELS

In this section, we consider the special case of polar coding for 3-user multiple access channels. Let W be a 3-user MAC with transition probability $W(y|x[1], x[2], x[3])$. The uniform rate region of W is in general a 3-dimensional polyhedron, which is shown for an example, described later, in Figure 3, where R_1, R_2 and R_3 are rates of user 1, 2 and 3, respectively.

Let $X[1], X[2]$ and $X[3]$ be uniform and independent binary random variables. Let I_1, I_2 and I_3 be defined as follows:

$$\begin{aligned} I_1 &= I(X[1]; Y) \\ I_2 &= I(X[2]; Y, X[1]) \\ I_3 &= I(X[3]; Y, X[1, 2]) \end{aligned}$$

The dominant face $\mathcal{D}(W)$ is the hexagon whose vertices are specified in Figure 1. These vertices are also called *corner points* of the uniform rate region. For simplicity, it is assumed that W is symmetric with respect to the inputs i.e. $W(y|x[1], x[2], x[3])$ remains the same if $(x[1], x[2], x[3])$ is permuted.

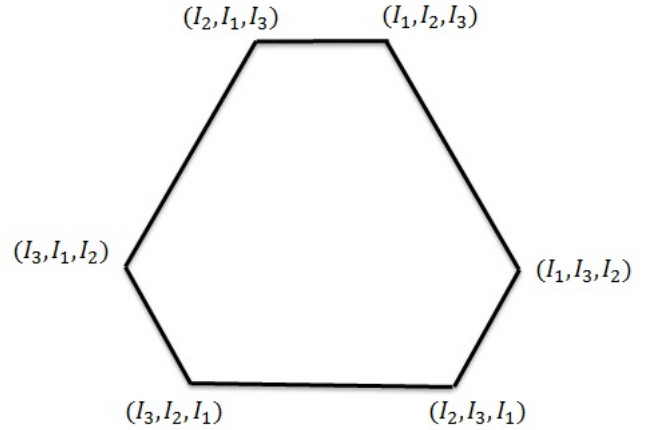


Fig. 1: The dominant face $\mathcal{D}(W)$

The corner points are achievable by *separate* polar coding. For instance, in order to achieve the point (I_1, I_2, I_3) , the message of user 1 is decoded first. Then the message of user 2 is decoded assuming the user 1's message is known. At the end, the message of user 3 is decoded, assuming the first and second messages are known. In this case, the scheme does not depend on the underlying polar codes and any other capacity achieving code will fit as well. We observed that the successive method of decoding the messages follows the same rule as in the successive cancellation decoding of polar codes. Therefore, we propose the joint successive cancellation decoding for all the users. The idea is to shuffle the decoding order of not only the messages, but also the bits within each message. In order to guarantee channel polarization, we let the decoding order for a certain polarization *base code* to be arbitrarily chosen. Then the recursive steps of channel polarization are applied on top of the base code.

For the separate polar coding, the length of the base code is 1, which is defined to be the length of the message for each user. There are $3! = 6$ different base codes which result

in the achievability of 6 corner points. The idea then is to increase the length of the base code and achieve more and more points. An example of a base code of length 2 is shown in Figure 2. The decoding order of this base code is given by $(x_1[1], x_1[2], x_1[3], x_2[1], x_2[2], x_2[3])$. This decoding order can be simply denoted by a permutation π which permutes the default vector $(x_1[1], x_2[1], x_1[2], x_2[2], x_3[1], x_3[2])$. In this case, $\pi = (1, 4, 2, 5, 3, 6)$. For polar coding with a general block length N built upon this base code, the decoding order is specified as follows. The successive cancellation decoding decodes the first half of user 1's message first, then the first half of user 2's message and then the first half of user 3's message followed by the second half of the messages of user 1, 2 and 3, respectively.

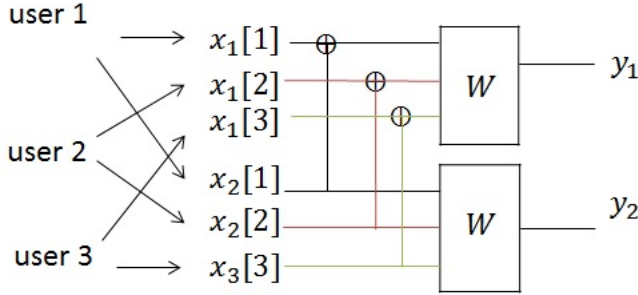


Fig. 2: An example for a base code of length 2

In total, there are $\frac{6!}{2!2!2!} = 90$ possible base codes of length 2. Notice that the base code has to keep the order within each user's message i.e. $x_1[j]$ has to appear before $x_2[j]$ in the decoding order. We have derived all these achievable points for an example. The system model assumes that the multiple access channel W is a binary-additive Gaussian noise channel where inputs $x[1], x[2], x[3] \in \{0, 1\}$ are modulated using BPSK (0 is mapped to -1 and 1 is mapped to $+1$) into $\bar{x}[1], \bar{x}[2]$ and $\bar{x}[3]$, respectively. The output of the channel is denoted by y , where $y = \bar{x}[1] + \bar{x}[2] + \bar{x}[3] + N$ and N is the Gaussian noise of unit variance N_0 . For this channel, the capacity region is same as the uniform rate region and is given by the set of all possible 3-tuple rates (R_1, R_2, R_3) that satisfy

$$\begin{aligned} R_1, R_2, R_3 &\leq I_3 = 0.7215 \\ R_1 + R_2, R_1 + R_3, R_2 + R_3 &\leq I_3 + I_2 = 1.1106 \\ R_1 + R_2 + R_3 &\leq I_1 + I_2 + I_3 = 1.3681 \end{aligned}$$

The uniform rate region along with all the 90 achievable points on the dominant face with base codes of length 2 are shown in Figure 3. We prove in the next section that by letting the length of the base code to grow large, the achievable 3-tuple rates of base codes cover the entire dominant face which will then be used to establish the achievability of the capacity region with polar coding.

IV. COVERING RADIUS OF THE DOMINANT FACE

Let W be a given m -user binary-input discrete multiple access channel. Let also $l \geq 1$ be a positive integer and $L = 2^l$. For $j = 1, 2, \dots, m$, assume that $U_1^L[j] =$

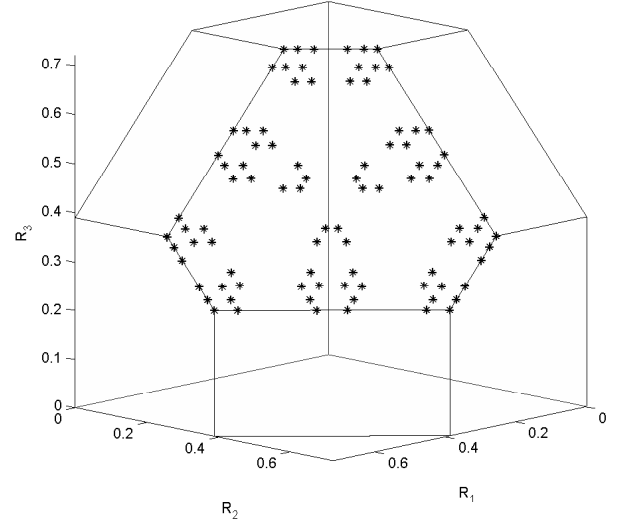


Fig. 3: The uniform rate region and the achievable points with base codes of length 2

$(U_1[j], U_2[j], \dots, U_L[j])$ is a vector of independent and uniformly distributed bits that is generated by the j -th user, independent of the other users. We also refer to $U_1^L[j]$ as the input bits of the j -th user. Let also $X_1^L[j] = U_1^L[j] G^{\otimes l}$. For $i = 1, 2, \dots, L$, the m -tuple $(X_i[1], X_i[2], \dots, X_i[m])$ is transmitted through the i -th independent copy of W and the output is denoted by Y_i . There are several ways to form an ordered sequence from the concatenated vector $(U_1^L[1], U_1^L[2], \dots, U_1^L[m])$. For notational convenience, the concatenated vector $(U_1^L[1], U_1^L[2], \dots, U_1^L[m])$ is also denoted by D_1^{mL} . In general, there are

$$\binom{mL}{L, L, \dots, L} = \frac{(mL)!}{L!^m}$$

permutations π assuming that the order of $U_1^L[j]$ is preserved through the permutation i.e. $U_i[j]$ appears in the permuted sequence before $U_k[j]$, for any j and $i < k$. Let \mathcal{P}_L denote the set of all such permutations. The described MAC polar transformation along with the choice of permutation $\pi \in \mathcal{P}_L$ represents a polarization base code of length L , on top of which the recursive channel polarization is applied. This procedure will be elaborated more in the next section. The permutation π will enforce the decoding order in the joint successive cancellation decoder and hence, is also referred to as the decoding order.

The mutual information $I(D_1^{mL}; Y_1^L)$ can be expanded with respect to the permutation π using the chain rule as follows:

$$I(D_1^{mL}; Y_1^L) = \sum_{i=1}^{mL} I(D_{\pi^{-1}(i)}; Y_1^L, D_{\pi^{-1}(1)}^{\pi^{-1}(i-1)})$$

Then for $i = 1, 2, \dots, mL$, let $I_i^{(\pi)}$ denote the i -th term in the above expansion i.e.

$$I_i^{(\pi)} \stackrel{\text{def}}{=} I(D_{\pi^{-1}(i)}; Y_1^L, D_{\pi^{-1}(1)}^{\pi^{-1}(i-1)}) \quad (8)$$

Also, for $j = 1, 2, \dots, m$, let

$$R_j^{(\pi)} \stackrel{\text{def}}{=} \frac{1}{L} \sum_{i=(j-1)L+1}^{jL} I_{\pi(i)}^{(\pi)} \quad (9)$$

In fact, intuitively speaking, $R_j^{(\pi)}$ is the allocated capacity to the j -th user. As we will see in the next section, the m -tuple of rates $\mathbf{R}^{(\pi)} = (R_1^{(\pi)}, R_2^{(\pi)}, \dots, R_m^{(\pi)})$ is achievable with polar coding built upon the base code with permutation π . But before that, through the rest of this section, we characterize the *covering radius* of the dominant face. The covering radius r , with respect to the length of the base codes L , is formally defined as follows:

$$r \stackrel{\text{def}}{=} \max_{\mathbf{Q} \in \mathcal{D}(W)} \min_{\pi \in \mathcal{P}_L} \|\mathbf{Q} - \mathbf{R}^{(\pi)}\| \quad (10)$$

where $\|\mathbf{X}\|$ is the Euclidean norm of \mathbf{X} in the m -dimensional space \mathbb{R}^m .

The following lemma is the result of (8) and (9).

Lemma 1: For any permutation π , the sum-rate of $\mathbf{R}^{(\pi)}$ is $I(W)$.

Remark. In fact a more general statement than Lemma 1 holds. For any permutation π , $\mathbf{R}^{(\pi)}$ is a point on the dominant face $\mathcal{D}(W)$. This will also follow as a result of the next section, where we prove that these points are achievable by polar coding.

For two users j_1 and j_2 , we write $j_1 \rightarrow j_2$ if an input bit of the user j_1 appears right before an input bit of the user j_2 through the permutation π . More precisely, there exist i_1 and i_2 with $(j_1 - 1)L + 1 \leq i_1 \leq j_1 L$ and $(j_2 - 1)L + 1 \leq i_2 \leq j_2 L$ such that $\pi(i_2) = \pi(i_1) + 1$. In that case, we also define the new permutation π' from π by swapping $\pi(i_1)$ and $\pi(i_2)$ i.e. $\pi'(i_2) = \pi(i_1)$, $\pi'(i_1) = \pi(i_2)$ and $\pi'(i) = \pi(i)$ for $i \neq i_1, i_2$. We say that π' is the transposition of π with respect to $j_1 \rightarrow j_2$. Notice that the transposition with respect to $j_1 \rightarrow j_2$ is not necessarily unique, as there may be other input bits of the user j_1 that appear right before other input bits of the user j_2 . In that case, we may choose one of them for the transposition.

Lemma 2: Let π' be the transposition of π with respect to $j_1 \rightarrow j_2$. Then we have

$$R_j^{(\pi)} = R_j^{(\pi')} \text{ for } j \neq j_1, j_2$$

and

$$0 \leq R_{j_1}^{(\pi')} - R_{j_1}^{(\pi)} = R_{j_2}^{(\pi)} - R_{j_2}^{(\pi')} \leq \frac{1}{L}$$

Proof: By definition, π' is the result of π by swapping $\pi(i_1)$ and $\pi(i_2)$, where $\pi(i_2) = \pi(i_1) + 1$. Therefore, by definition of $I_i^{(\pi)}$ in (8), $I_i^{(\pi)} = I_i^{(\pi')}$ for $i \neq i_1, i_2$. This implies that the values of $R_j^{(\pi)}$ do not change during the specified transposition, for $j \neq j_1, j_2$.

For the second part,

$$\begin{aligned} L(R_{j_1}^{(\pi')} - R_{j_1}^{(\pi)}) &= \frac{\sum_{i=(j_1-1)L+1}^{j_1 L} I_i^{(\pi')}}{L} - \frac{\sum_{i=(j_1-1)L+1}^{j_1 L} I_i^{(\pi)}}{L} \\ &= \frac{I_{i_1}^{(\pi')} - I_{i_1}^{(\pi)}}{L} \end{aligned} \quad (11)$$

Also,

$$0 \leq I_{i_1}^{(\pi)} \leq I_{i_1}^{(\pi')} \leq 1 \quad (12)$$

(11) and (12) together imply that

$$0 \leq R_{j_1}^{(\pi')} - R_{j_1}^{(\pi)} \leq \frac{1}{L}$$

The other inequality also follows similarly or simply by observing that $\sum R_j^{(\pi)} = \sum R_j^{(\pi')} = I(W)$, by Lemma 1. ■

Fix an arbitrary point $\mathbf{Q} = (Q_1, Q_2, \dots, Q_m)$ on the dominant face $\mathcal{D}(W)$. For a given permutation π , let $\mathcal{A}_Q^{(\pi)} \subset [m]$ be defined as follows:

$$\mathcal{A}_Q^{(\pi)} \stackrel{\text{def}}{=} \{j \in [m] : R_j^{(\pi)} < Q_j\} \quad (13)$$

Lemma 3: For any $\mathcal{B} \subset \mathcal{A}_Q^{(\pi)}$, at least one input bit from the complement set \mathcal{B}^c appears after some input bits of the set \mathcal{B} through the permutation π .

Proof: Assume, to the contrary, that all input bits of the users in the set \mathcal{B}^c appear before all the input bits of the users in the set \mathcal{B} . This together with the chain rule of mutual information and definition of $R_j^{(\pi)}$ in (9) imply that

$$\sum_{j \in \mathcal{B}} R_j^{(\pi)} = I(U_1^L[\mathcal{B}]; Y, U_1^L[\mathcal{B}^c]). \quad (14)$$

Notice that \mathbf{Q} is a point included in $\mathcal{S}(W)$. Therefore, using (7) and (14) we get

$$\sum_{j \in \mathcal{B}} Q_j \leq I(U_1^L[\mathcal{B}]; Y, U_1^L[\mathcal{B}^c]) = \sum_{j \in \mathcal{B}} R_j^{(\pi)}$$

But since $\mathcal{B} \subset \mathcal{A}_Q^{(\pi)}$, for any $j \in \mathcal{B}$, $Q_j > R_j^{(\pi)}$ which is a contradiction. This proves the lemma. ■

For two users j and j' , we say that j' is reachable from j , if there exists a sequence of users j_1, j_2, \dots, j_t , with $j_i \in [m]$, such that $j \rightarrow j_0 \rightarrow j_1 \rightarrow \dots \rightarrow j_t \rightarrow j'$. This sequence is also referred to as the path from j to j' . If such path exists, then one can assume, without loss of generality, that j_1, j_2, \dots, j_t are distinct. Clearly, if j'' is reachable from j' and j' is reachable from j , then j'' is reachable from j .

Corollary 4: For any $j \in \mathcal{A}_Q^{(\pi)}$, at least one element of the complement set $[m] - \mathcal{A}_Q^{(\pi)}$ is reachable from j .

Proof: Let \mathcal{B} be the set of all reachable users from j . Then \mathcal{B} is a closed set in the sense that no element in \mathcal{B}^c is reachable from any element of \mathcal{B} . It implies that through the permutation π , all the input bits of \mathcal{B} appear after all the input bits of \mathcal{B}^c . Then by Lemma 3, \mathcal{B} is not a subset of $\mathcal{A}_Q^{(\pi)}$ which proves the corollary. ■

Now that we have established the necessary notations and derived desired properties of the m -tuple rates of the base codes, we turn to state the main theorem of this section as follows:

Theorem 5: Given a point $\mathbf{Q} \in \mathcal{D}(W)$, there exists a permutation $\pi \in \mathcal{P}_L$ such that for any index $j \in [m]$, we have

$$|Q_j - R_j^{(\pi)}| \leq \frac{m-1}{L}$$

Proof: Let π to be the permutation such that the Euclidean distance $\|\mathbf{R}^{(\pi)} - \mathbf{Q}\|$ is minimum among all the permutations in \mathcal{P}_L i.e. for any other permutation π' :

$$\|\mathbf{R}^{(\pi)} - \mathbf{Q}\| \leq \|\mathbf{R}^{(\pi')} - \mathbf{Q}\|.$$

Consider $\mathcal{A}_Q^{(\pi)}$, defined as in (13). If $\mathcal{A}_Q^{(\pi)}$ is empty, then it implies that for any $j \in [m]$, $Q_j \leq R_j^{(\pi)}$. But both $\mathbf{R}^{(\pi)}$ and \mathbf{Q} have the same sum-rate i.e.

$$\sum_{j \in [m]} R_j^{(\pi)} = \sum_{j \in [m]} Q_j = I(W).$$

Therefore, $\mathbf{Q} = \mathbf{R}^{(\pi)}$ and the lemma is proved. In the case that $\mathcal{A}_Q^{(\pi)}$ is non-empty, let j_1 be an arbitrary element of $\mathcal{A}_Q^{(\pi)}$. By Corollary 4, there is a path $j_1 \rightarrow j_2 \rightarrow \dots \rightarrow j_t$ with $j_i \in \mathcal{A}_Q^{(\pi)}$, for $i = 2, \dots, t-1$, and $j_t \notin \mathcal{A}_Q^{(\pi)}$. Without loss of generality, we can assume that j_i 's are distinct and hence $t \leq m$. For notational convenience, let's re-label the users so that the user j_i is labeled with i for $i \in [t]$.

Define the sequence of real numbers $\{a_i\}_{i=1}^t$, where $a_i = Q_i - R_i^{(\pi)}$. We claim that $a_{i-1} - a_i \leq \frac{1}{L}$. The claim will imply that $a_1 - a_t \leq \frac{t-1}{L}$ and since a_t is not a positive number, as $t \notin \mathcal{A}_Q^{(\pi)}$, we will have

$$a_1 = Q_1 - R_1^{(\pi)} \leq a_1 - a_t \leq \frac{t-1}{L} \leq \frac{m-1}{L}$$

By symmetry, the same set of arguments can be applied to the set $[m] - \mathcal{A}_Q^{(\pi)}$, where one can prove that $R_j^{(\pi)} - Q_j \leq \frac{m-1}{L}$, for $j \in [m] - \mathcal{A}_Q^{(\pi)}$. This will complete the proof of theorem. What remains is to prove the claim for the sequence $\{a_i\}_{i=1}^t$.

For $i = 2, 3, \dots, t$, let the permutation π_i be the transposition of π with respect to $i-1 \rightarrow i$. By Lemma 2, we have

$$0 \leq R_{i-1}^{(\pi_i)} - R_{i-1}^{(\pi)} = R_i^{(\pi)} - R_i^{(\pi_i)} \leq \frac{1}{L} \quad (15)$$

$$R_j^{(\pi)} = R_j^{(\pi_i)} \text{ for } j \neq i-1, i \quad (16)$$

On the other hand, by the choice of π , we know that

$$\|\mathbf{R}^{(\pi)} - \mathbf{Q}\| \leq \|\mathbf{R}^{(\pi_i)} - \mathbf{Q}\| \quad (17)$$

(16) and (17) together imply that

$$\begin{aligned} (Q_{i-1} - R_{i-1}^{(\pi)})^2 + (Q_i - R_i^{(\pi)})^2 \\ \leq (Q_{i-1} - R_{i-1}^{(\pi_i)})^2 + (Q_i - R_i^{(\pi_i)})^2 \end{aligned} \quad (18)$$

Let $\alpha = R_{i-1}^{(\pi_i)} - R_{i-1}^{(\pi)}$, then by (15), (18) can be re-written as

$$a_{i-1}^2 + a_i^2 \leq (a_{i-1} - \alpha)^2 + (a_i + \alpha)^2$$

which can be simplified as

$$a_{i-1} - a_i \leq \alpha$$

By (15), $\alpha \leq \frac{1}{L}$ which completes the proof of claim. ■

Corollary 6: The covering radius of the dominant face $\mathcal{D}(W)$ is upper bounded by $\frac{(m-1)\sqrt{m}}{L}$.

Proof: Consider an arbitrary point \mathbf{Q} on the dominant face. Then the permutation π exists as in Theorem 5. Then we will have

$$\|\mathbf{Q} - \mathbf{R}^{(\pi)}\|^2 = \sum_{j=1}^m (Q_j - R_j^{(\pi)})^2 \leq \frac{m(m-1)^2}{L^2}.$$

which proves the corollary by definition of covering radius in (10). ■

Remark. In Theorem 5 of [11], we showed for a two-user MAC that for any point \mathbf{Q} on the dominant face there exists a pair of achievable rate with a distance at most $\frac{\sqrt{2}}{L}$ from \mathbf{Q} . In other words, the covering radius for the special case of $m = 2$ is upper bounded by $\frac{\sqrt{2}}{L}$. This result matches with the result of Corollary 6 for $m = 2$.

V. ACHIEVING THE UNIFORM RATE REGION WITH JOINT DECODING

In this section, we show that the points $\mathbf{R}^{(\pi)}$ defined in (9) are achievable by polar coding and a joint successive cancellation decoder, for all $L = 2^l$ and permutations $\pi \in \mathcal{P}_L$.

Let W be a given m -user binary-input discrete multiple access channel. Let also $n \geq l$ be a positive integer and $N = 2^n$. For $j = 1, 2, \dots, m$, assume that $U_1^N[j]$ is a vector of independent and uniformly distributed bits that is generated by the j -th user, independent of other users. Let also $X_i^N[j] = U_1^N[j] G^{\otimes n}$. For $i = 1, 2, \dots, N$, the m -tuple $(X_i[1], X_i[2], \dots, X_i[m])$ is transmitted through the i -th independent copy of W and the output is denoted by Y_i . This transformation together with an ordered sequence of the input bits $(U_1^N[1], U_1^N[2], \dots, U_1^N[m])$ is called a MAC polar transformation. This ordered sequence is also referred to as the decoding order, because it will specify the order in which input bits are decoded in the successive cancellation decoder. For a fixed value of L , we limit our attention to polarization transforms of length N whose decoding orders are built upon elements of \mathcal{P}_L , specified next.

For any permutation $\pi : [M] \rightarrow [M]$ and any $k \in \mathbb{N}$, where $k, M \in \mathbb{N}$, we define the permutation $\pi^{(k)} : [kM] \rightarrow [kM]$, as follows. First the sequence $(1, 2, \dots, kM)$ is split into k sub-sequences s_1, s_2, \dots, s_M of length k each, where $s_i = (k(i-1) + 1, k(i-1) + 2, \dots, ki)$. Then the permutation π is applied to the sequence (s_1, s_2, \dots, s_M) to get the sequence $s_{\pi(1)}, s_{\pi(2)}, \dots, s_{\pi(M)}$. At the end, the sub-sequences are expanded back to get a sequence of total length kM . For instance, if $M = 2$ and $\pi(1) = 2, \pi(2) = 1$, then for any k , the permutation $\pi^{(2k)}$ replaces the first and second sub-block of length k with each other.

With the above definition, decoding orders on $(U_1^N[1], U_1^N[2], \dots, U_1^N[m])$ specified by $\pi^{(N/L)}$, for some $\pi \in \mathcal{P}_L$, are considered. In that sense, the MAC polar transformations of length L and decoding orders π are the corresponding base codes. As we will see, the polar transformations of length N built upon polarization base codes of length L can be decomposed into mL certain single-user polar transformations of length N/L .

Let $W^N : \mathcal{X}^{mN} \rightarrow \mathcal{Y}^N$ denote the channel consisting of N independent copies of W i.e.

$$W^N(y_1^N | x_1^N[\mathcal{M}]) \stackrel{\text{def}}{=} \prod_{i=1}^N W(y_i | x_i[\mathcal{M}]) \quad (19)$$

where $\mathcal{M} = [m]$ for notational convenience. The combined channel \widetilde{W}_N is defined with transition probabilities given by

$$\widetilde{W}_N(y_1^N | u_1^N[\mathcal{M}]) \stackrel{\text{def}}{=} W^N(y_1^N | x_1^N[\mathcal{M}]) \quad (20)$$

where $x_1^N[j] = u_1^N[j] \cdot G^{\otimes n}$, for $j \in \mathcal{M}$. Let d_1^{mN} denote the ordered version of the sequence $(u_1^N[1], u_1^N[2], \dots, u_1^N[m])$ according to a certain $\pi^{(N/L)}$. The bit-channels are defined with respect to this ordered sequence, which will also enforce the decoding order in the joint successive decoding. For $j = 1, 2, \dots, mN$, the j -th bit-channel is defined as

$$W_N^{(j)}(y_1^N, d_1^{j-1} | d_j) \stackrel{\text{def}}{=} \frac{1}{2^{mN-1}} \sum_{d_{j+1}^{mN} \in \{0,1\}^{mN-j}} \widetilde{W}_N(y_1^N | d_1^{mN}) \quad (21)$$

The permutation π and consequently the decoding order are assumed to be fixed for the rest of this section. Actually the results are not specific to π and will hold for any $\pi \in \mathcal{P}_L$.

The following lemma is similar to the Proposition 3 of [1] that still holds for the particular MAC polar transformation defined here and its corresponding bit-channels.

Lemma 7: Given $N = 2^n \geq L$ independent copies of W and the bit-channels defined in (21), for any j with $1 \leq j \leq N$,

$$W_{2N}^{(2j-1)} = W_N^{(j)} \boxtimes W_N^{(j)}$$

and

$$W_{2N}^{(2j)} = W_N^{(j)} \otimes W_N^{(j)}$$

Proof: Let $d_{1,o}^{2N}$ and $d_{1,e}^{2N}$ denote the even-indexed and odd-indexed subvectors, respectively. Then

$$\begin{aligned} W_{2N}^{(2i)} &= \frac{1}{2^{2mN-1}} \sum_{d_{2i+1}^{2mN}} \widetilde{W}_N(y_1^N | d_{1,o}^{2N} \oplus d_{1,e}^{2N}) \widetilde{W}_N(y_1^N | d_{1,e}^{2N}) \\ &= \frac{1}{2} \frac{1}{2^{mN-1}} \sum_{d_{2i+1,e}^{2mN}} \widetilde{W}_N(y_1^N | d_{1,e}^{2N}). \\ &\quad \frac{1}{2^{mN-1}} \sum_{d_{2i+1,o}^{2mN}} \widetilde{W}_N(y_1^N | d_{1,o}^{2N} \oplus d_{1,e}^{2N}) \\ &= W_N^{(i)} \otimes W_N^{(i)} \end{aligned}$$

where we used the definitions of bit-channels in (21) and the channel combining operation in (2) together with the recursive structure of polar transformation. The other equation can be derived using the similar arguments. ■

Lemma 7 implies that the polar transformation of length N and decoding order $\pi^{(N/L)}$ can be regarded as $n - l$ levels of polarization of the bit-channels defined with polarization base code of length L and decoding order π . This is the result of following corollary:

Corollary 8: For a given $k \in [mL]$, let \overline{W} denote the single user channel $W_L^{(k)}$. Let also $n \geq l$ and $N = 2^n$. Then for any $\frac{(k-1)N}{L} + 1 \leq j \leq \frac{kN}{L}$,

$$W_N^{(j)} = \overline{W}^{(j - \frac{(k-1)N}{L})}$$

where the bit-channels with respect to \overline{W} are defined as in (5).

Proof: The proof is by induction on n . The base of induction is trivial for $n = 1$. The induction steps are by the fact that channel combining recursion steps in Arikan's polar transformation, as proved in Proposition 3 of [1], match with Lemma 7. ■

For $j \in [m]$, let $S_L^{(j)}$ denote the set of positions of the input bits of the j -th user in the ordered sequence specified by π .

Then the set of indices $[mN]$ is split into m subsets $S_N^{(j)}$, for $j \in [m]$, as follows:

$$S_N^{(j)} = \left\{ j \in [mN] : \left\lfloor \frac{Lj}{N} \right\rfloor \in S_L^{(j)} \right\}$$

In fact, $S_N^{(j)}$ is the set of positions assigned to the j -th user in the decoding order specified by $\pi^{(N/L)}$. The set of good bit-channels for the j -th user is also defined as follows. For any $\beta < \frac{1}{2}$ and $N = 2^n$

$$\mathcal{G}_N^{(j)}(W, \beta) \stackrel{\text{def}}{=} \left\{ j \in S_N^{(j)} : Z(W_N^{(j)}) < 2^{-N^\beta}/mN \right\} \quad (22)$$

The next theorem establishes the main result of channel polarization for the MAC polar transformation.

Theorem 9: For any m -user binary-input discrete MAC W , any constant $\beta < \frac{1}{2}$ and $j \in [m]$, we have

$$\lim_{N \rightarrow \infty} \frac{|\mathcal{G}_N^{(j)}(W, \beta)|}{N} = R_j^{(\pi)}$$

Proof: By Corollary 8,

$$\mathcal{G}_N^{(j)}(W, \beta) = \bigcup_{i \in S_L^{(j)}} \mathcal{G}_{\frac{N}{L}}^{(i)}(W_L^{(i)}, \beta)$$

There exists $\beta' < \frac{1}{2}$ such that $N^\beta < (\frac{N}{L})^{\beta'}$ for large enough N . Therefore, channel polarization theorem for the single user case [13] imply that

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{|\mathcal{G}_N^{(j)}(W, \beta)|}{N} &\geq \frac{1}{L} \sum_{i \in S_L^{(j)}} \lim_{N \rightarrow \infty} \frac{|\mathcal{G}_{\frac{N}{L}}^{(i)}(W_L^{(i)}, \beta)|}{N/L} \\ &= \frac{1}{L} \sum_{i \in S_L^{(j)}} I_i^{(\pi)} = R_j^{(\pi)} \end{aligned}$$

Since the fraction of all good bit-channels can not exceed the sum-rate $I(W) = \sum_{j \in [m]} R_j^{(\pi)}$, both the above inequalities should be equality. This completes the proof of theorem. ■

A. Encoding, decoding and asymptotic performance

The encoding of the proposed scheme in the previous subsection is similar to that of original polar codes for each of the users. In fact, regardless of the choice of polarization base code the encoder for any block length N is fixed, where each user multiplies a vector of length N by the polarization matrix $G^{\otimes n}$. The only thing that depends on the polarization base code is the indices of information bits. For a fixed length of the polarization base code $L = 2^l$ and decoding order π , the set of good bit-channels for all the m users $\mathcal{G}_N^{(j)}(W, \beta)$ are defined in (22). In the polar encoder for the j -th user, u_i is an information bit for any $i \in \mathcal{G}_N^{(j)}(W, \beta)$. Otherwise, u_i is frozen to a fixed value. Therefore, the rate of polar code for the j -th user approaches $R_j^{(\pi)}$, as N goes to infinity, by Theorem 9. The frozen bits are chosen uniformly at random, revealed to the receiver a priori and fixed during the course of communication. If the underlying channels are symmetric, then the frozen bits can be set to zeros.

For joint decoding, the successive cancellation decoding defined in [1] is performed, with some modification. The order of decoding is determined by $\pi^{(N/L)}$, defined in the previous

subsection. The low-complex implementation of successive cancellation decoder invented by Arkan in [1] can be extended to our scheme for MAC. The recursive steps of computing likelihood ratios can be done in a similar way using mL trellises of size $N \times (n - l + 1)$, corresponding to the mL bit-channels defined with respect to the polarization base code of length L . The input likelihood ratios (LR) for these trellises can be computed using a naive way with complexity $O(2^{mL})$ for each LR. Notice that this is the worst case required complexity in the sense that depending on the choice of particular base code one can invoke the recursive structure, imposed by $G^{\otimes l}$ for each of the users, to compute the LRs more efficiently. The total decoding complexity is then upper bounded by $O(mN(n - l + 1 + 2^{mL}))$, where m and L are regarded as fixed parameters in the scheme. Therefore, the decoding complexity is asymptotically $O(N \log N)$, similar to the original polar decoding.

For the asymptotic performance of the proposed polar coding, the following theorem follows similar to [1] and [13].

Theorem 10: For any $\beta < \frac{1}{2}$, any m -user MAC W , any $\epsilon > 0$ and any point \mathbf{Q} on $\mathcal{D}(W)$, there exists a family of polar codes that approaches a point on the dominant face within distance ϵ from \mathbf{Q} . Furthermore, the probability of frame error under successive cancellation decoding is less than 2^{-N^β} , where N is the block length of the code for each of the users.

Proof: The choice of L for the polarization base code depends on ϵ . Fix $L = 2^l$ such that $\frac{(m-1)\sqrt{m}}{L} < \epsilon$. Then by Corollary 6 there exists π such that the distance between \mathbf{Q} and $\mathbf{R}^{(\pi)}$ is less than ϵ . Fixing L and π , for any block length $N = 2^n \geq L$, we construct the polar code for the j -th user with respect to the set of good bit-channels $\mathcal{G}_N^{(j)}(W, \beta)$ defined in (22). Then by Theorem 9, the m -tuple of rates approach $\mathbf{R}^{(\pi)}$ as N goes to infinity. The probability of frame error is bounded by the sum of the Bhattacharyya parameters of the selected bit-channels. Therefore, it is less than 2^{-N^β} for the joint successive cancellation decoding of all the m messages together, by definition of $\mathcal{G}_N^{(j)}(W, \beta)$. ■

B. Comparison with prior works

Our approach for polarizing MAC is essentially different from those of [5], [6]. The authors of [5], [6] propose a framework for MAC polarization by extending the notion of channel splitting from the single user case to the two user case [5], and then to the m -user case [6], while we propose a method to view the MAC polarization as a single user channel polarization by considering all the possible decoding orders. First, we compare the two schemes in terms of decoding complexity and then capacity achieving property.

Both of our scheme and those of [5], [6] use a successive cancellation decoding, originally proposed by Arkan in [1]. As discussed in the foregoing subsection, we have to combine the low-complex Arkan's decoder with a basic decoder for the base code. As a result, we have an extra additive term $O(mN2^{mL})$ in the complexity which is dominated by $O(mN(n - l + 1))$, as N goes to infinity, while L and m are assumed to be constant. On the other hand, in the scheme proposed in [5] and extended in [6], the likelihood of a vector of length m needs

to be tracked along the decoding trellis. Therefore, instead of a simple likelihood ratio, a vector of length 2^m for the probability of all 2^m possible cases has to be computed recursively. This increases the decoding complexity by a factor of 2^m . In fact, the decoding complexity is still $O(N \log N)$, but if we look at the actual number of operations needed to complete the decoding, the decoding complexity of the scheme in [5], [6] is $\frac{2^m}{m}$ times more than the decoding complexity of our scheme, asymptotically.

Next, we compare the two schemes in terms of capacity-achieving property. In Theorem 10, we proved that all the points on the dominant face of the uniform rate region can be achieved with arbitrary small resolution. As pointed out before, the scheme proposed in [5], [6] does not necessarily achieve the whole uniform rate region. It is only guaranteed that one point on the dominant face is achievable. Here, we provide an example for a two user MAC such that the scheme proposed in [5], achieves only one point on the dominant face. The example can be extended to the case that the number of users is a power of two.

Let $W' : \{0, 1\} \rightarrow \mathcal{Y}'$ be an arbitrary single user binary-input DMC. Then the two user multiple access channel $W : \{0, 1\} \rightarrow \mathcal{Y}$ is constructed as follows. Let $\mathcal{Y} = \mathcal{Y}' \times \mathcal{Y}'$. For any two binary inputs u and v corresponding to the first and second user, $u \oplus v$ and v are transmitted through two independent copies of W' with outputs y_1 and y_2 . Then the output of W is defined to be $y = (y_1, y_2) \in \mathcal{Y}$. The diagram of W is depicted in Figure 4.

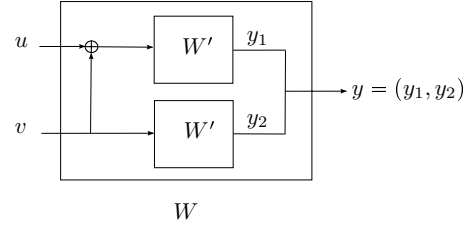


Fig. 4: A constructed example for two user MAC

Let $I = I(W')$, $I^- = I(W'^-)$ and $I^+ = I(W'^+)$. Then the uniform rate region of W is as shown in Figure 5.

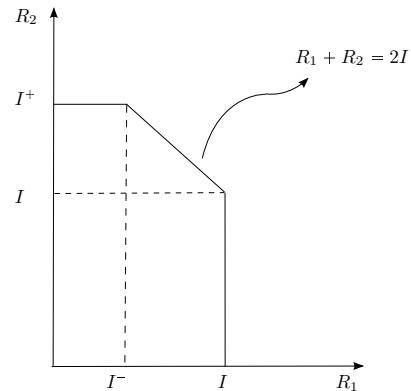


Fig. 5: The uniform rate region for the constructed example

It is easy to observe that a polar transformation of length N on two user MAC W , in the sense defined in [5], is equivalent

to a single user polar transformation of length $2N$ over W' . Let U_1^N and V_1^N be the input bits by the first and the second user, respectively. Let $W_N^{(i)}$ be the i -th two user bit-channel observed by (U_i, V_i) . Then

$$I^{(2)}(W_N^{(i)}) = I(V_i; Y_1^i, U_1^i, V_1^{i-1}) = I(W_{2N}^{\prime(2i)})$$

Also,

$$\begin{aligned} I^{(1)}(W_N^{(i)}) &= I(U_i; Y_1^i, U_1^{i-1}, V_1^i) \\ &\geq I(U_i; Y_1^i, U_1^{i-1}, V_1^{i-1}) = I(W_{2N}^{\prime(2i-1)}) \end{aligned}$$

And,

$$I(W_N^{(i)}) = I(W_{2N}^{\prime(2i)}) + I(W_{2N}^{\prime(2i-1)})$$

Notice that as N goes to infinity, $W_{2N}^{\prime(2i-1)}$ and $W_{2N}^{\prime(2i)}$ are either both good or both not good, for almost all i 's. In fact, the fraction of i 's for which only one of them is good approaches zero. Therefore, the triple $(I^{(1)}(W_N^{(i)}), I^{(2)}(W_N^{(i)}), I(W_N^{(i)}))$ approaches either $(0, 0, 0)$ or $(1, 1, 2)$ among the five possible cases discussed in [5]. The point $(0, 0, 0)$ corresponds to the case that both of the inputs needs to be frozen to fixed values. The point $(1, 1, 2)$ corresponds to the case that both the inputs carry information bits. Thus, the only achievable point on the dominant face by this scheme is the point (I, I) , and the achievable region is the square with vertices $(0, 0)$, $(0, I)$, $(I, 0)$, (I, I) . On the other hand, as proved in this section, the whole region shown in Figure 5 can be achieved by our proposed scheme with joint successive decoder.

VI. DISCUSSIONS AND CONCLUSION

In this paper, we considered the problem of designing polar codes for transmission over general m -user multiple access channels. The key observation behind our work is to view the polar transformations for the m -users across the multiple access channel as a unified MAC polar transformation. We showed that the MAC polar transformation with a certain decoding order can be split into mL single user polar transformations, where L is the length of the base code. We also proved that the covering radius of the dominant face is upper bounded by $\frac{(m-1)\sqrt{m}}{L}$. Therefore, by letting L to grow large, we are able to achieve the entire uniform rate region with the constructed MAC polar code.

The idea behind the joint successive cancellation decoding is that decoding the messages of different users successively is essentially following the same rule as in the successive cancellation decoder of polar codes. Therefore, one can invoke the joint successive cancellation decoder with a certain decoding order that is the result of shuffling the input bits of all users. We also derived the direct relation between the decoding order and the polarization base codes in order to guarantee channel polarization. We proved that by shuffling the order of the information bits of different users to be decoded, which will also enforce the choice of MAC polarization base code, the achievability of the entire uniform rate region is guaranteed.

Since the individual codes for each of the users can be regarded as single user polar codes, all the existing methods for improving the performance of single user polar codes can be applied on top of our proposed MAC-polar codes. For instance,

the individual polar codes can be made systematic as suggested by Arikan [16] in order to improve the bit error rate. Also, concatenation of polar codes with other block codes [17] or list decoding of polar codes [18] can be used to boost the finite-length performance. Also, as discussed in [11], one can invoke the compound polar codes proposed in [19] in order to improve the finite length performance if time sharing is used.

REFERENCES

- [1] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3051–3073, 2009.
- [2] H. Mahdaviyar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6428–6443, 2011.
- [3] E. Arikan, "Source polarization," *Proceedings of IEEE International Symposium on Information Theory*, pp. 899–903, 2010.
- [4] E. Abbe, "Randomness and dependencies extraction via polarization," *Proceedings of Information Theory and Applications Workshop (ITA)*, pp. 1–7, 2011.
- [5] E. Şaşoğlu, E. Telatar, and E. Yeh, "Polar codes for the two-user binary-input multiple-access channel," *Proceedings of IEEE Information Theory Workshop (ITW)*, pp. 1–5, 2010.
- [6] E. Abbe and E. Telatar, "Polar codes for the m -user multiple access channel," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5437–5448, 2012.
- [7] R. Ahlswede, "Multi-way communication channels," *Proceedings of 2nd International Symposium on Information Theory*, Tsahkadsor, Armenia SSR, 1971.
- [8] H. H.-J. Liao, "Multiple access channels," Ph.D. dissertation, Univ. Hawaii, Honolulu, 1972.
- [9] E. Arikan, "Polar coding for the slepian-wolf problem based on monotone chain rules," *Proceedings of IEEE International Symposium on Information Theory*, pp. 566–570, 2012.
- [10] H. Mahdaviyar, M. El-Khamy, J. Lee, and I. Kang, "Achieving the uniform rate region of multiple access channels using polar codes," *arXiv preprint arXiv:1307.2889*, July 10, 2013.
- [11] —, "Techniques for polar coding over multiple access channels," *Proceedings of 48th Annual Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, USA, March 19–21, 2014.
- [12] S. Onay, "Successive cancellation decoding of polar codes for the two-user binary-input mac," *Proceedings of IEEE International Symposium on Information Theory*, pp. 1122–1126, July 7–12, 2013.
- [13] E. Arikan and E. Telatar, "On the rate of channel polarization," *Proceedings of IEEE International Symposium on Information Theory*, pp. 1493–1495, 2009.
- [14] S. B. Korada, "Polar codes for channel and source coding," Ph.D. dissertation, École Polytechnique Fédérale De Lausanne, 2009.
- [15] S. B. Korada, E. Şaşoğlu, and R. Urbanke, "Polar codes: Characterization of exponent, bounds, and constructions," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 6253–6264, 2010.
- [16] E. Arikan, "Systematic polar coding," *IEEE communications letters*, vol. 15, no. 8, pp. 860–862, 2011.
- [17] H. Mahdaviyar, M. El-Khamy, J. Lee, and I. Kang, "Performance limits and practical decoding of interleaved reed-solomon polar concatenated codes," *IEEE Transactions on Communications*, vol. 62, no. 5, pp. 1406–1417, 2014.
- [18] I. Tal and A. Vardy, "List decoding of polar codes," *Proceedings of IEEE International Symposium on Information Theory*, pp. 1–5, 2011.
- [19] H. Mahdaviyar, M. El-Khamy, J. Lee, and I. Kang, "Compound polar codes," *Proceedings of Information Theory and Applications Workshop (ITA)*, pp. 1–6, San Diego, California, 2013.